UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 16669.12-M
2. GOVT ACCESSION NO.: AD-A109 132
3. RECIPIENT'S CATALOG NUMBER:

4. TITLE (and Subtitle): Comparison of the Bioptimal Curve with Curves for Two Robust Estimates
5. TYPE OF REPORT & PERIOD COVERED: Technical
6. PERFORMING ORG. REPORT NUMBER:

7. AUTHOR(s): Katherine Bell Krystinik, Stephan Morgenthaler
8. CONTRACT OR GRANT NUMBER(s): DAAG29 79 C 0205

9. PERFORMING ORGANIZATION NAME AND ADDRESS: Princeton University, Princeton, NJ 08544
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS:

11. CONTROLLING OFFICE NAME AND ADDRESS: U. S. Army Research Office, Post Office Box 12211, Research Triangle Park, NC 27709
12. REPORT DATE: Oct 81
13. NUMBER OF PAGES: 11

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):
15. SECURITY CLASS. (of this report): Unclassified
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

NA

18. SUPPLEMENTARY NOTES

The view, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

curves (mathematics)
estimating
Gaussian functions
weighting

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The performance of one-step biweights and Bob Bell estimates are compared with the performance of the bioptimal estimates. This is done for sample sizes 5, 10, and 20. The two situations are the Gaussian and slash.

8112 31087

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED 4/10 A 73
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Comparison of the Bioptimal Curve with Curves for Two Robust Estimates*

by

Katherine Bell Krystinik
and
Stephan Morgenthaler

Technical Report No. 195, Series 2
Department of Statistics
Princeton University
October 1981

(Corrected)

*Prepared in connection with research at Princeton University, supported by the Army Research Office (Durham).

ABSTRACT

The performance of one-step biweights and Bob Bell estimates are compared with the performance of the bioptimal estimates. This is done for sample sizes 5,10, and 20. The two situations are the Gaussian and slash.

Note: This version has been corrected, both as plots and tables. Copies of the uncorrected version should be destroyed.

This is a revised version. The numbers and figures in the old version were wrong due to wrong weighting.

Bioptimal estimates (with respect to to the slash and Gaussian) are estimates whose performance cannot be improved simultaneously in both situations. Performance is measured in efficiencies where minimum attainable single situation variances are obtained via configural polysampling (Pregibon and Tukey (1981), Bell and Pregibon (1981)). Each bioptimal estimate corresponds to a ratio of shadow prices. For a given pair of shadow prices and a specific data configuration the bioptimal estimate is a simple weighted mean of the two single situation optimal estimates for that configuration (Krystinik (1981)). (This is true for any other invariant estimates except for cases, of total probability zero, where the two optimal estimators coincide.) Averaging over configurations yields efficiencies (Pregibon and Tukey (1981)). These points form the bioptimal curve shown in figure 1 (2,3) for sample size 5 (10,20).

The bioptimal curves are compared to the curves (for varying c values) of the one-step biweight (starting point the median and denominator c·MAD) and the estimates proposed by B. Bell with ψ function:

Prepared in part in connection with research at Princeton University, supported by the Army Research Office (Durham).

October 27, 1981

$$\psi(x) = x(1 + \frac{x^2}{3})^{-2}$$

Where $x = (y-T)/c \cdot MAD$ and $T_0 = med$ (Bell (1980)). Curves for the one-step version of this estimate are shown in figures 1-3. Points corresponding to the iterated version have also been plotted.

The horizontal lines in the figures show where the 95% and 90% Gaussian-efficient estimates lie on the curve. The bioptimal estimate which lies on the diagonal is the minimax estimate for the Gaussian and slash and the criterion

$$\frac{\text{variance for a situation}}{\text{minimum variance for that situation}} .$$

The point labelled with "bieffective" corresponds to the minimax estimate conditioned on the configuration (using the weights of the configuration).

Tables 1 (2,3) show the variances in the two situations, Gaussian and slash, corresponding to the points in figures 1 (2,3).

REFERENCES

Bell, B. (1980). "Adaptive scale estimation in the location problem," Technical Report, Department of Statistics, Stanford University.

Bell, K. and Pregibon, D.P. (1981). "Some computational details of configural sampling methods, Technical Report No. 191, Series 2, Department of Statistics, Princeton University.

Krystinik, Katherine Bell (1981). "Data modifications based on order: Pushback; a configural polysampling approach," Ph.D. Thesis, Department of Statistics Princeton University, Princeton, New Jersey.

Pregibon, D. and Tukey, J.W. (1981). "Assessing the behavior of robust estimates of location in small samples: Introduction to configural polysampling," Technical Report No. 185, Series 2, Department of Statistics, Princeton University.

Figure 1

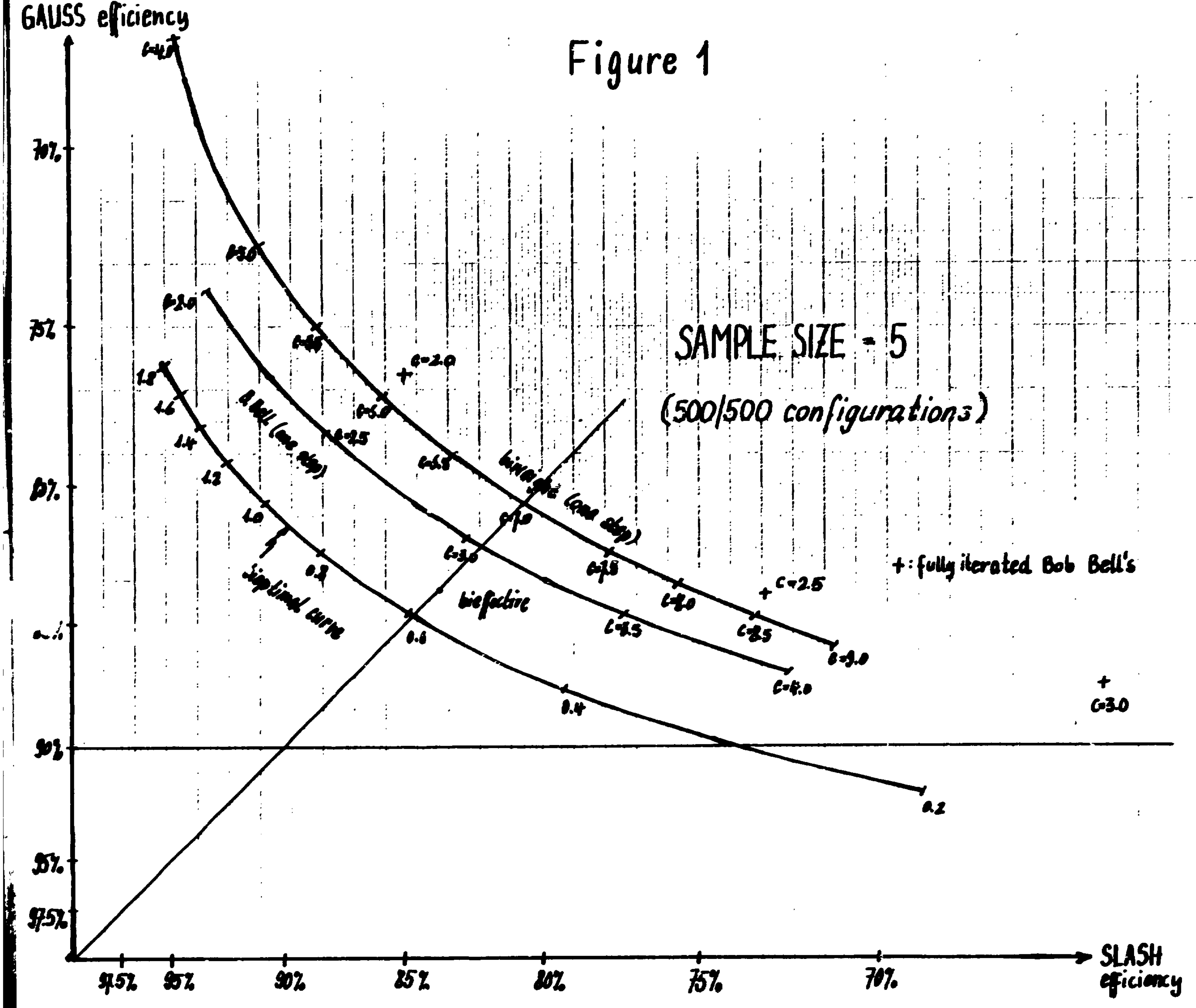
GAUSS efficiency
SAMPLE SIZE = 5
(500/500 configurations)
+ : fully iterated Bob Bell's
biefactive
Sigmoidal curve
SLASH efficiency
97.5%
95%
90%
85%
80%
75%
70%

Figure 2

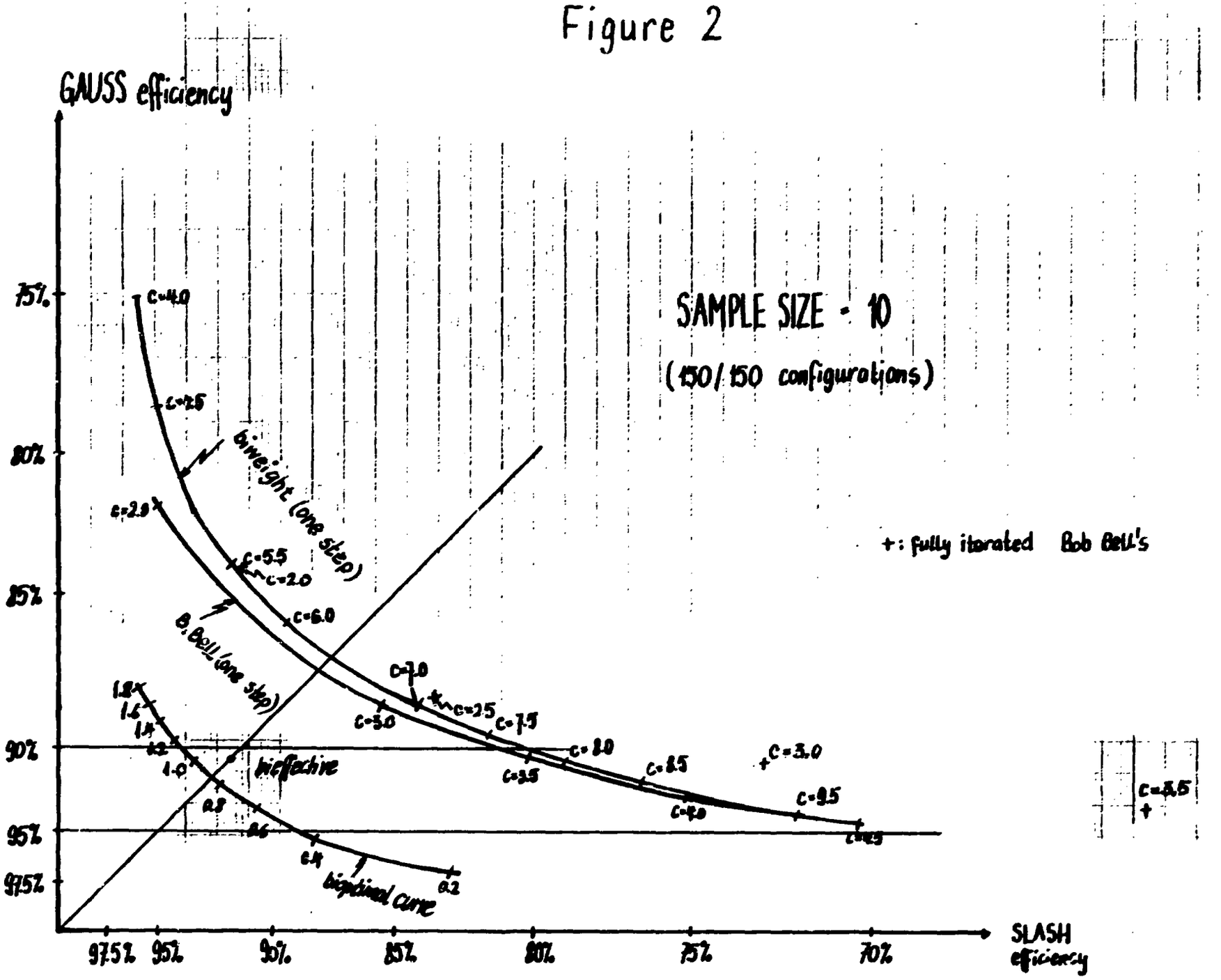
GAUSS efficiency
75%
80%
85%
90%
95%
97.5%
SAMPLE SIZE - 10
(150/150 configurations)
+: fully iterated Bob Bell's
c=4.0
c=2.5
biweight (one step)
c=2.0
c=5.5
c=6.0
B. Bell (one step)
c=3.0
c=7.0
+c=2.5
c=7.5
c=8.0
c=8.5
c=9.0
c=9.5
+c=3.0
+c=5.5
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
biefficient
bisquare curve
97.5%
95%
90%
85%
80%
75%
70%
SLASH efficiency

Figure 3

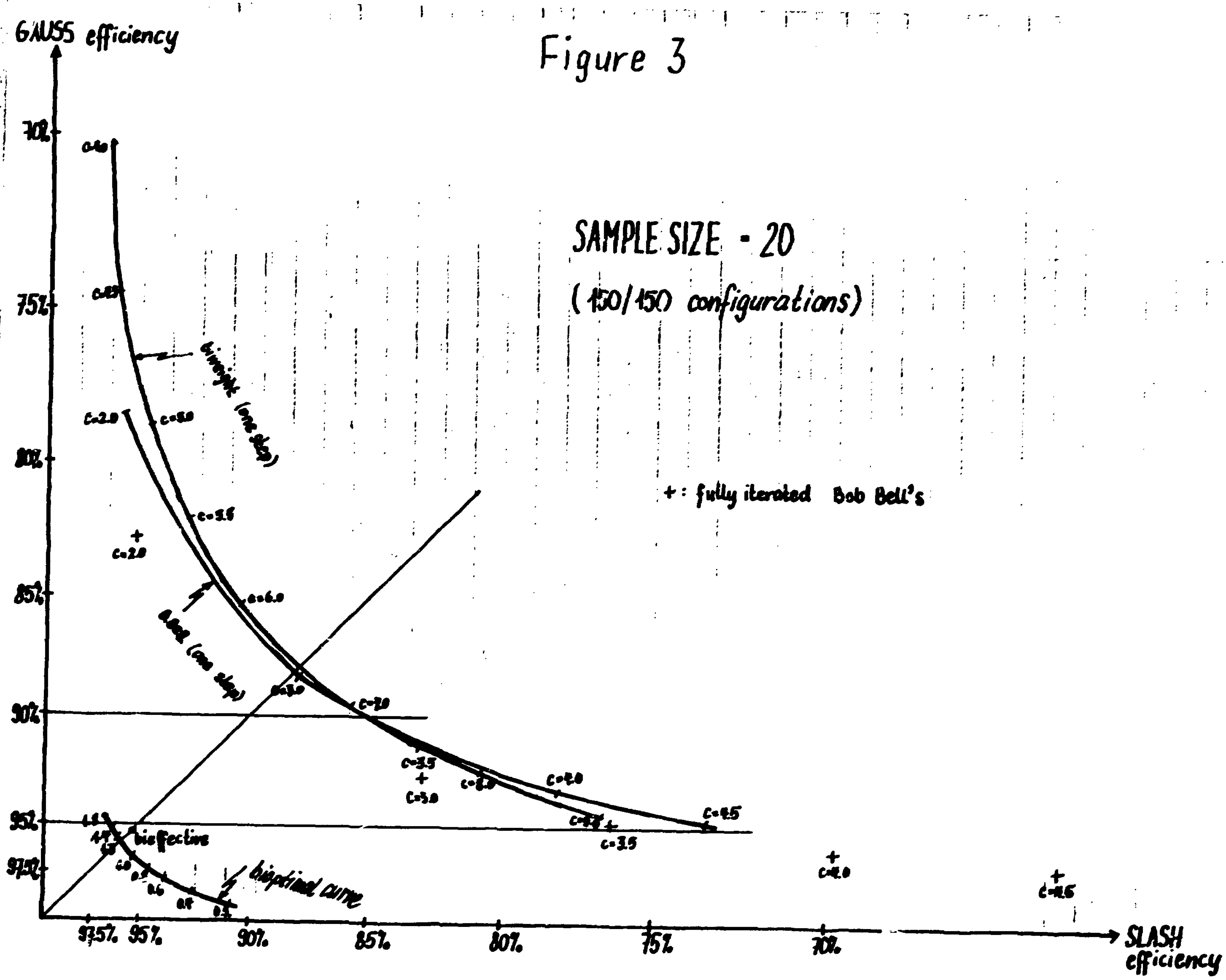
GAUSS efficiency
SAMPLE SIZE - 20
(150/150 configurations)
+ : fully iterated Bob Bell's
biweight (one step)
biefective
bioptimal curve
70%
75%
80%
85%
90%
95%
97.5%
SLASH efficiency

Table 1

Variances n=5 (estimated by configural polysampling)

optimal variances Gaussian: 0.2012* slash: 1.7904

bioptimal curve shadow price ratio (S/G)	Gaussian	slash
1.8	.2540	1.8784
1.4	.2579	1.9128
1.2	.2541	1.9388
1.0	.2497	1.9747
0.8	.2444	2.0274
0.6	.2380	2.1104
0.4	.2298	2.2587
0.2	.2189	2.5997
biefficient 0.59	.2377	2.1146
one-step biweights constant		
4.0	.2991	1.8890
5.0	.2769	1.9618
5.5	.2683	2.0227
6.0	.2610	2.0847
6.5	.2548	2.1513
7.0	.2495	2.2209
7.5	.2450	2.2927
8.0	.2411	2.3662
8.5	.2377	2.4472
9.0	.2347	2.5175
one-step Bob Bell constant		
2.0	.2722	1.9216
2.5	.2570	2.0306
3.0	.2460	2.1650
3.5	.2379	2.3143
4.0	.2318	2.4726
iterated Bob Bell constant		
2.0	.2632	2.1047
2.5	.2400	2.4485
3.0	.2303	2.7717
3.5	.2245	3.0758
4.0	.2195	3.4757
bieffective	.2405	2.1404

*Theoretical value .2000. Correlated sampling and matching cubature ensures that efficiency comparisons with values in the column below are relatively more precise. All values given are for the weighted sample of configurations analyzed;divison by .2012/.2000 should improve essentially every number in this column if absolute, rather than relative, values are needed.

Table 2

Variances n=10 (estimated by configural polysampling)

optimal variances	Gaussian: 0.0936*	slash: 0.6354
bioptimal curve shadow price ratio (S/G)	Gaussian	slash
1.8	.1056	.6627
1.2	.1031	.6744
0.8	.1010	.6891
0.6	.0997	.7015
0.4	.0984	.7219
0.2	.0964	.7662
biefficient 0.847	.1013	.6869
one-step biweights constant		
4.0	.1247	.6619
4.5	.1194	.6689
5.5	.1116	.6940
6.0	.1089	.7120
7.0	.1048	.7552
7.5	.1033	.7795
8.0	.1020	.8049
8.5	.1010	.8306
9.5	.0994	.8831
one-step Bob Bell constant		
2.0	.1145	.6690
3.0	.1049	.7431
3.5	.1022	.7921
4.0	.1003	.8461
4.5	.0990	0.9036
iterated Bob Bell constant		
2.0	.1115	.6966
2.5	.1052	.7614
3.0	.1019	.8720
3.5	.0997	.9999
4.0	.0978	1.2521
bieffective	.1021	.6926

*Theoretical value .1000. Correlated sampling and matching cubature ensures that efficiency comparisons with values in the column below are relatively more precise. All values given are for the weighted sample of configurations analyzed; division by .0936/.1000 should improve essentially all numbers in this column if absolute, rather than relative, values are needed.

Table 3

Variances n=20 (estimated by configural polysampling)

optimal variances	Gaussian: .0528*	slash: 0.2534
bioptimal curve shadow price ratio (S/G)	Gaussian	slash
1.8	.0557	.2622
1.4	.0552	.2637
1.0	.0546	.2659
0.8	.0543	.2676
0.6	.0540	.2698
0.4	.0536	.2731
0.2	.0533	.2794
biefficient 1.3	.0550	.2642
one-step biweights constant		
4.0	.0751	.2611
4.5	.0708	.2632
5.0	.0672	.2674
5.5	.0643	.2729
6.0	.0621	.2795
7.0	.0595	.2959
8.0	.0571	.3138
9.0	.0559	.3310
one-step Bob Bell constant		
2.0	.0674	.2639
3.0	.0599	.2886
3.5	.0579	.3058
4.0	.0566	.3248
4.5	.0557	.3452
iterated Bob Bell constant		
2.0	.0638	.2655
3.0	.0569	.3058
3.5	.0556	.3319
4.0	.0549	.3631
4.5	.0544	.3928
bieffective	.0553	.2655

*Theoretical value .0500. Correlated sampling and matching cubature ensures that efficiency comparisons with values in the column below are relatively more precise. All values given are for the weighted sample of configurations analyzed; division by .0500/.0528 should improve essentially all numbers in this column if absolute, rather than relative, values are needed.